African Journal of Agricultural Research Vol. 7(14), pp. 2096-2113, 12 April, 2012
Available online at http://www.academicjournals.org/AJAR
DOI: 10.5897/AJAR11.627
ISSN 1991-637X ©2012 Academic Journals

*Full Length Research Paper*

# A study on the west Sahel rainfall variability: The role of the intertropical convergence zone (ITCZ)

**P. S. Lucio[1,2]*, L. C. B. Molion[4], F. C. Conde[3] and M. L. D. de Melo[4]**

[1]Centro de Geofísica de Évora (CGE) Apartado 94, 7000-554 Évora, Portugal.
[2]Universidade Federal do Rio Grande do Norte (UFRN), Centro de Ciências Exatas e da Terra (CCET), Departamento de Estatística 59078-970 Natal-RN, Brazil.
[3]Instituto Nacional de Meteorologia (INMET) Brasília-DF, Brazil.
[4]Universidade Federal de Alagoas (UFAL) Maceió-AL, Brazil.

Accepted 24 August, 2011

**This article provides a report of rainfall analysis results supporting the hypothesis that rainfall variability in the Western Sahel is associate with temporal variability of the incursions of ITCZ over West Africa. It is pointed out that the most significant climatic rainfall changes in Sahel most probably occurred between 1950 and 1975 with the decrease of the annual rainfall totals, a high reduction (about 70%) over the whole region. One of the most significant climatic variations has been the rainfall persistent decline in the Sahel since late 1960's. Remarkable latitudinal shift of ITF mean position towards the South seems to generate an overall reduction of annual rainfall. In the last two decades covered in this research (1985–2004), rainfall showed some improvement, particularly in the 1990–1999 decade. However, the more humid conditions of the 1950's and 1960's were not reestablished yet. The trend was abruptly interrupted by a return of adequate rainfall conditions in 1994. This was considered to be the wettest of the past 30 years and was thought to indicate the end of the drought perhaps. However, the 1994 rainfall total barely exceeded the long-term mean. The 1994 rainy season was unusual in the sense that the anomalously wet conditions occurred in the end of the rainy season and the months following. The temporal characteristics of the series, such as variance, were evaluated using principal components regression.**

**Key words:** Climate variability, cluster analysis, discriminant analysis, principal components analysis, principal component regression, trend analysis.

## INTRODUCTION

Africa is a vast continent, and it experiences a wide variety of climate regimes. The location, size, and shape of the African continent play key roles in determining climate. The poleward extremes of the continent experience winter rainfall associated with the passage of mid-latitude air masses. Across the Kalahari and Sahara deserts, precipitation is inhibited by subsidence virtually throughout the year. In contrast, moderate to heavy precipitation associated with the intertropical Convergence Zone (ITCZ) characterizes equatorial and tropical areas. Because the movement of the ITCZ follows the position of maximum surface heating associated with meridional displacement of the overhead position of the sun, near-equatorial regions experience two rain seasons, whereas regions further poleward experience one distinct rainfall season. The mean climate of Africa is further modified by the presence of large contrasts in topography and the existence of large lakes in some parts of the continent.

The Sahel region is the semi-arid belt in Africa between the arid to hyper-arid Sahara Desert and the humid equatorial Africa (Figure 1). In the 1970's and 1980's it experienced repeated, devastating droughts, most notably in 1972 to 1973 and in 1982 to 1984. Since then, scientists have been debating whether the cause for this pronounced climatic shift was anthropogenic in nature,

*Corresponding author. E-mail: pslucio@ccet.ufrn.br, pslucio@uevora.pt.

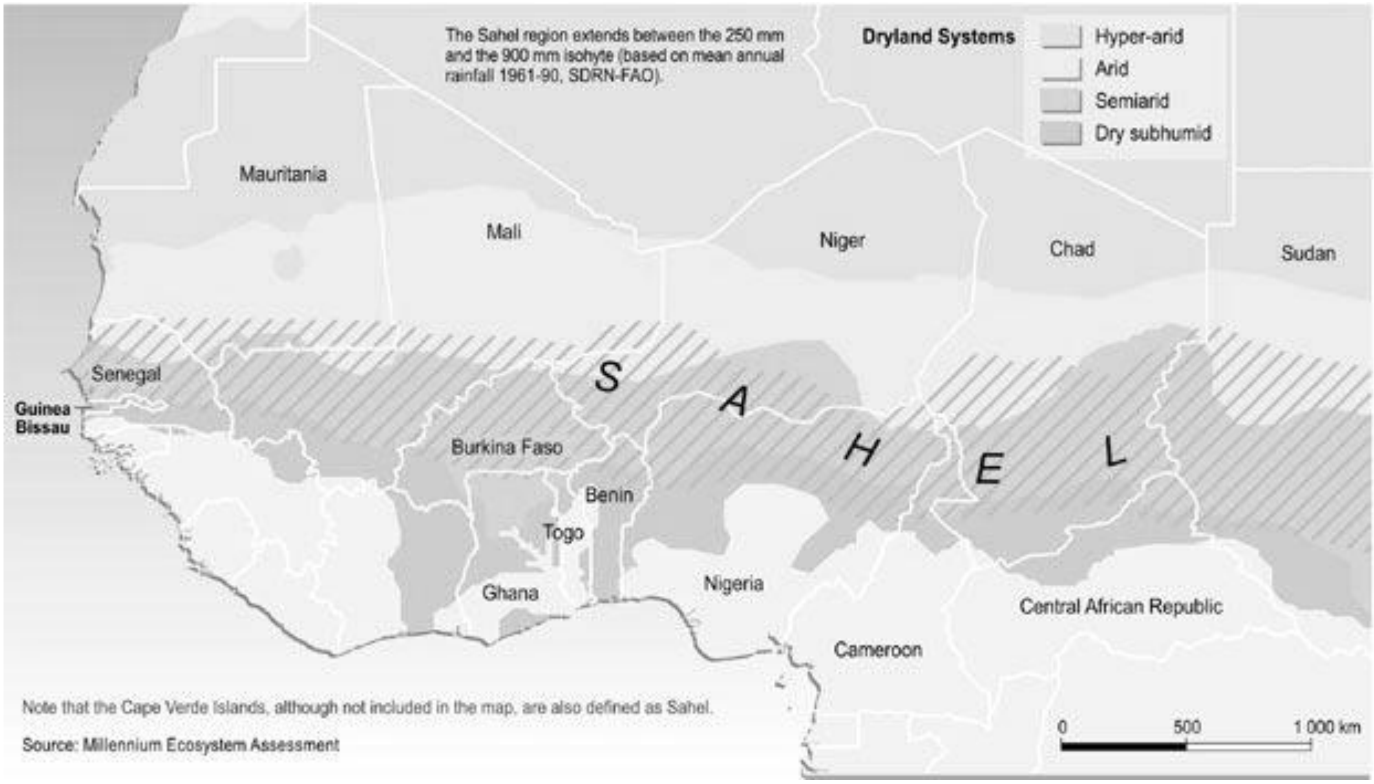


**Figure 1.** The West Sahel Africa region.

that is, whether desertification, forced by the increasing pressure of population growth on the environment, was starting to have an impact on the regional climate (Charney et al., 1975; Charney, 1975), or whether the recurrent droughts were related to the slower, decadal time scales of oceanic variability (Folland et al, 1986).

The drying trend in the semiarid Sahel was attributed to warmer-than-average low-latitude waters around Africa (Giannini et al., 2003), which, by favoring the establishment of deep convection over the ocean, weaken the continental convergence associated with the monsoon and engender widespread drought over the West Africa. Effectively, the relationship between rainfall and desert margin advances or retreats brings into focus questions regarding the interactions of global change dynamics with desertification processes. The fluctuation in the size of the Sahara Desert shows a linear inverse relationship between rainfall and expansion of the desert's southern margin (Milich, 1997). There is a vast literature available on this subject.

This manuscript analyzes the annual and seasonal rainfall variability across the West Sahel zone from 14 synoptic stations from 1950 to 2004, in view of numerous statistical techniques used to interpret spatial and monthly variability in West Sahel, suggesting (via statistical evidence) that variability of rainfall in the Sahel results from the temporal variability of the southern incursions of ITCZ on West Africa.

## OBSERVATIONAL DATASET

In this study, for the analysis and characterization of the precipitation system, the observed data of the World Monthly Surface Station Climatology (WMSSC) were used, giving larger focus to the periods with observed northward shift in the Sahara/Sahel boundary (Harrison et al., 1998). The dataset is supplied by the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) that fosters research collaboration between NOAA and the University of Washington in the period between 1950 to 2004 for 14 stations of the 6 countries of the area of West Africa Sahel (see the political boundaries and the precipitation climatology in Figure 1, whose coordinates are represented in the Table 1).

The period from 1950 to 2004 was analyzed – 1950 to 2000 from the JISAO dataset and 2001 to 2004 from the National Climatic Data Center (NCDC) dataset. This period was chosen due to well-documented shifts in rainfall climate on a decadal timescale in the West African region during this time. To deal with missing data problem we applied a multiple imputation technique. The Multivariate Imputation by Chained Equations (MICE) from R (R Development Core Team, 2009) was the software employed for imputing incomplete multivariate data. MICE is an R package implementing multiple imputation of incomplete multivariate data according the principle of Fully Conditional Specification (FCS). MICE consists of a set of flexible tools for creating multiple imputations and for the analysis of multiply imputed data sets (Rubin, 1987; Rubin, 1996).

## SEASONAL VARIABILITY ANALYSIS

It is arguable that the Sahel is the only one of the world's dry land regions to have experienced a significant

**Table 1.** The rain gauge West Sahel Africa stations (1950-2000).

| WMO ID | | STATION LOCATION |
|---|---|---|
| 1 | 610360 | TILLABERY - Niger (14.2°N, 1.5°E) |
| 2 | 610430 | TAHOUA - Niger (14.9°N, 5.3°E) |
| 3 | 610800 | MARADI - Niger (13.5°N, 7.1°E) |
| 4 | 612260 | GAO - Mali (16.3°N, 0.1°W) |
| 5 | 612700 | KITA - Mali (13.1°N, 9.5°W) |
| 6 | 612720 | SEGOU - Mali (13.4°N, 6.2°W) |
| 7 | 612930 | KOUTIALA - Mali (12.4°N, 5.5°W) |
| 8 | 614980 | KIFFA - Mauritania (16.6°N, 11.4°W) |
| 9 | 616410 | DAKAR - Senegal (14.7°N, 17.5°W) |
| 10 | 616540 | THIES - Senegal (14.8°N, 17.0°W) |
| 11 | 616660 | DIOURBEL - Senegal (14.8°N, 16.3°W) |
| 12 | 653060 | KANDI - Benin (11.1°N, 2.9°E) |
| 13 | 653190 | NATITINGUE - Benin (10.3°N, 1.4°E) |
| 14 | 655020 | OUGIHOUA – Burkina Faso (13.6°N, 2.4°W) |

The Time Series Data Credit: JISAO Climate Data Archive
http://tao.atmos.washington.edu/data_sets/sahel/

change[1] in climate resulting in increased aridity this century. Many authors have described the period since the late 1960's as representing a desiccation, that is, a phase characterized by persistent drought. Hulme (1996) describes this recent desiccation as contributing to an equivalent mean linear trend amounting to a 21% decrease in rainfall over the twentieth century. The downward trend in rainfall does not begin at the start of the dry episode; aggregated Sahel rainfall systematically declines from 1954. Dry episodes occurred in the 1910's and 1940's, but were of relatively short duration. The 1930's and 1950's were dominated almost entirely by wet years, with no significant rainfall deficits (there were no rainfall deficit years in the 1950's and there are no statistical significance indicating a behavior below average). 1950 and 1954 were the wettest years in the 20th century.

It should be noted that the rainfall deficits discussed in this manuscript represent annual rainfall amounts spatially averaged over the entire Sahelian region. Such aggregated values do not tell us anything about variations in the spatial distribution or rainfall seasonality. The latter can be as important as total rainfall amount in terms of agriculture. Agnew and Chappell (2000) point out that not all regions of the Sahel have experienced a synchronous and systematic reduction in rainfall, and that agricultural production in the Sahelian nations has increased over the course of the period generally characterized as dry. Indeed, studies have identified different regions of coherent rainfall (seasonal variability) within the Sahelian zone, and different areas are described by different, and sometimes conflicting, anomaly patterns. The ensemble displayed by Figure 2 illustrates the seasonal analysis with charts of seasonal indices and the Box-Whiskers plots of the monthly rainfall decomposition to study the nature of the component parts. As expected, the residuals (assigned as "random") are proportional to the monthly mean rainfall contribution. The additive model assumes that as the data increase the trend and the seasonal components are considered independents and then added to the random error component. There is a strong empirical evidence of seasonality, defining three composed-periods (precipitation regimes) by: dry, transition and wet. The August variability has an important contribution for the rainy periods. The dry regime corresponds to the months from November to May; the transition regime was associated for the months of June and October; and the rainy regime for the months of July, August and September. The rainy regime corresponds to the period the ITCZ reaches its more northern position in its movement to the North and the activity of a classical SW monsoon is more intense. The oscillatory movement of the ITCZ is a signature of the same phenomenon while in the dry regime; the ITCZ migrates to the South. Over the transition regime one still has some influence of the ITCZ and of the trade winds. It is well-known that western Africa is influenced by centers of high subtropical pressures of the Azores (in the North Atlantic region) and the Santa Helena (in the South Atlantic region), which control the seasonal oscillation of the trade winds with marine influence; the marine trade winds, and of continental characteristic, the continental trade winds. In the dry regime, the trade winds of NE are more intense. During the rainy regime, they are more parallel to the coast, becoming marine trade winds. Another prominence factor is the oscillatory movements of the ITCZ, configured by the migration of the high subtropical

[1] The current version is MICE V2.7, which is described in Van Buuren S, Groothuis-Oudshoorn K (2011). MICE: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software. More information and papers about MICE can be found in http://www.multiple-imputation.com/.

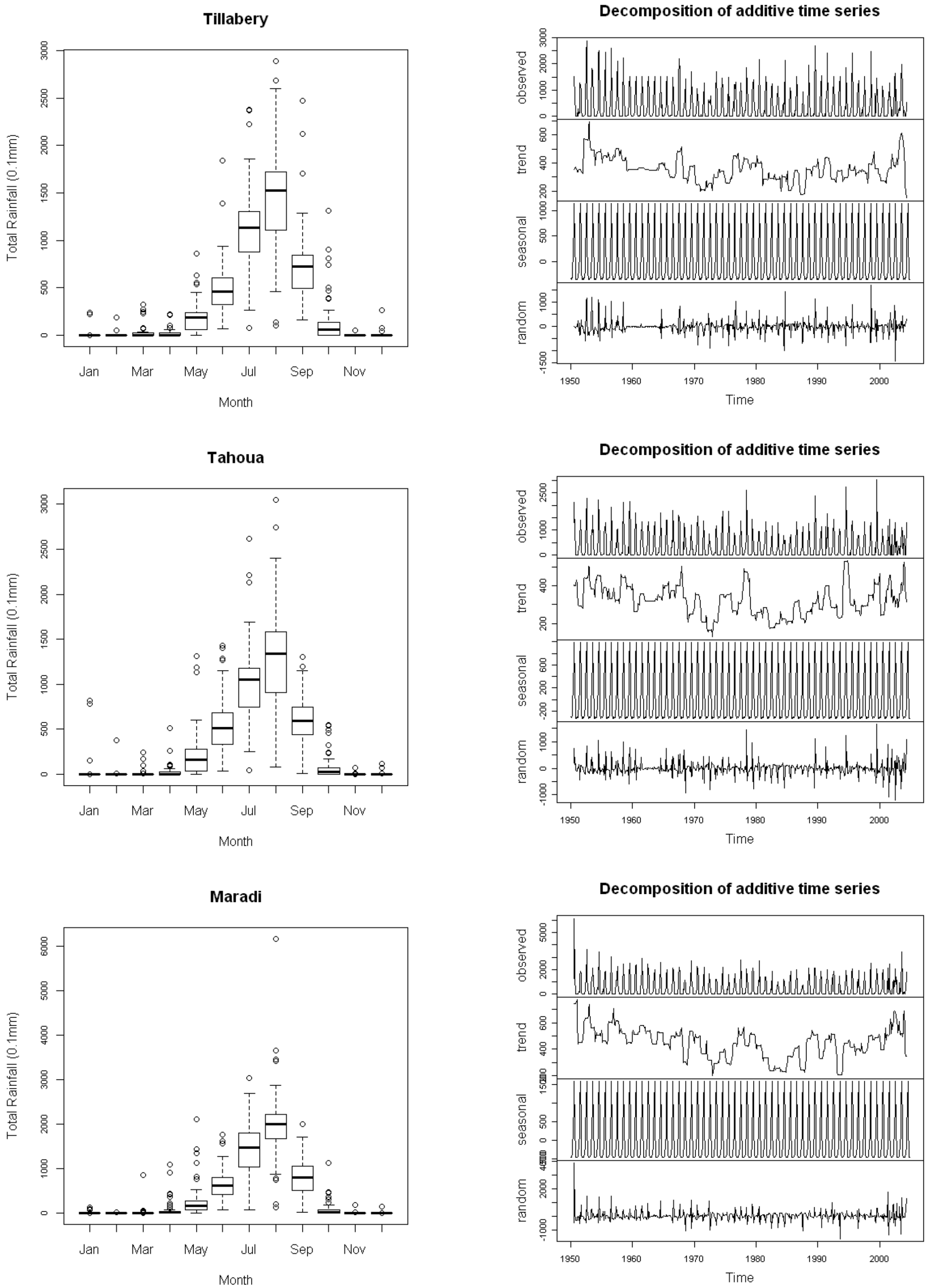


**Figure 2.** Left: The box-Whiskers plots of the rainfall data for each raingauge station. Right: The charts of seasonal indices and trend detection. Plots are presented of monthly rainfall decomposition for the climate period (1950 to 2000), separating each observed meteorological time series into seasonal, trend and random components. In the additive model, the trend and the seasonal components are considered independents and then added to the random error component.

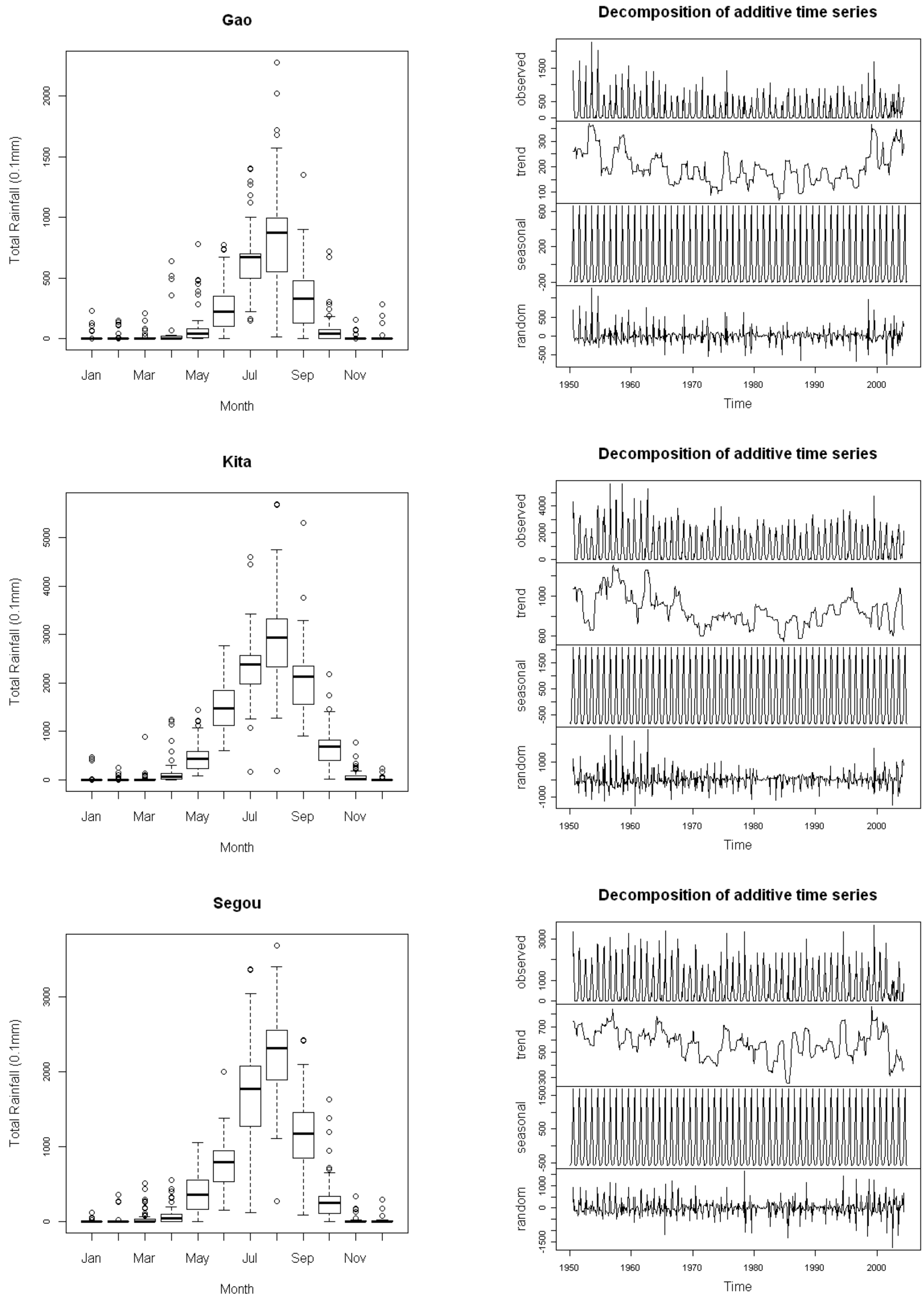


**Figure 2.** Contd.

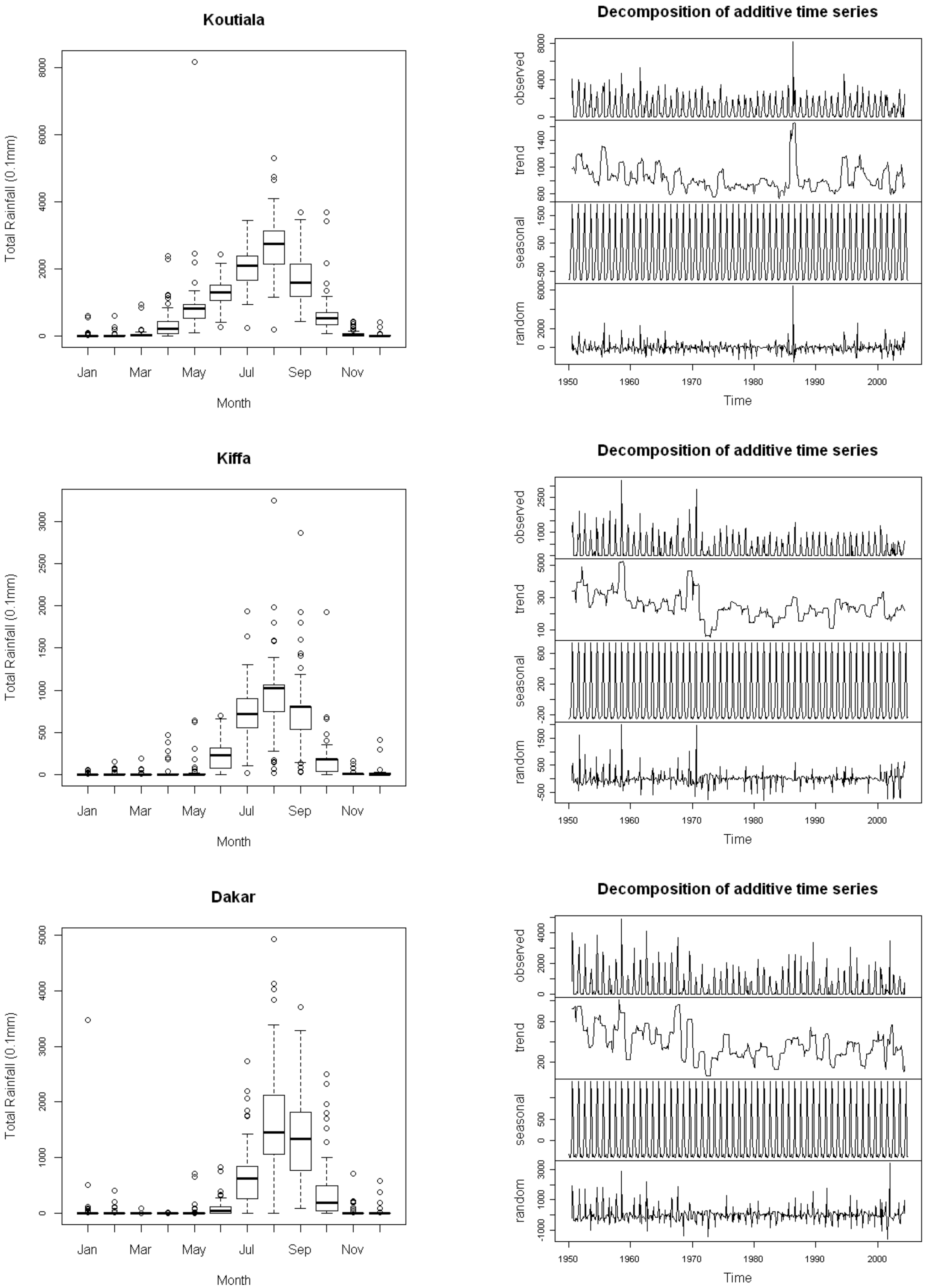


**Figure 2.** Contd.

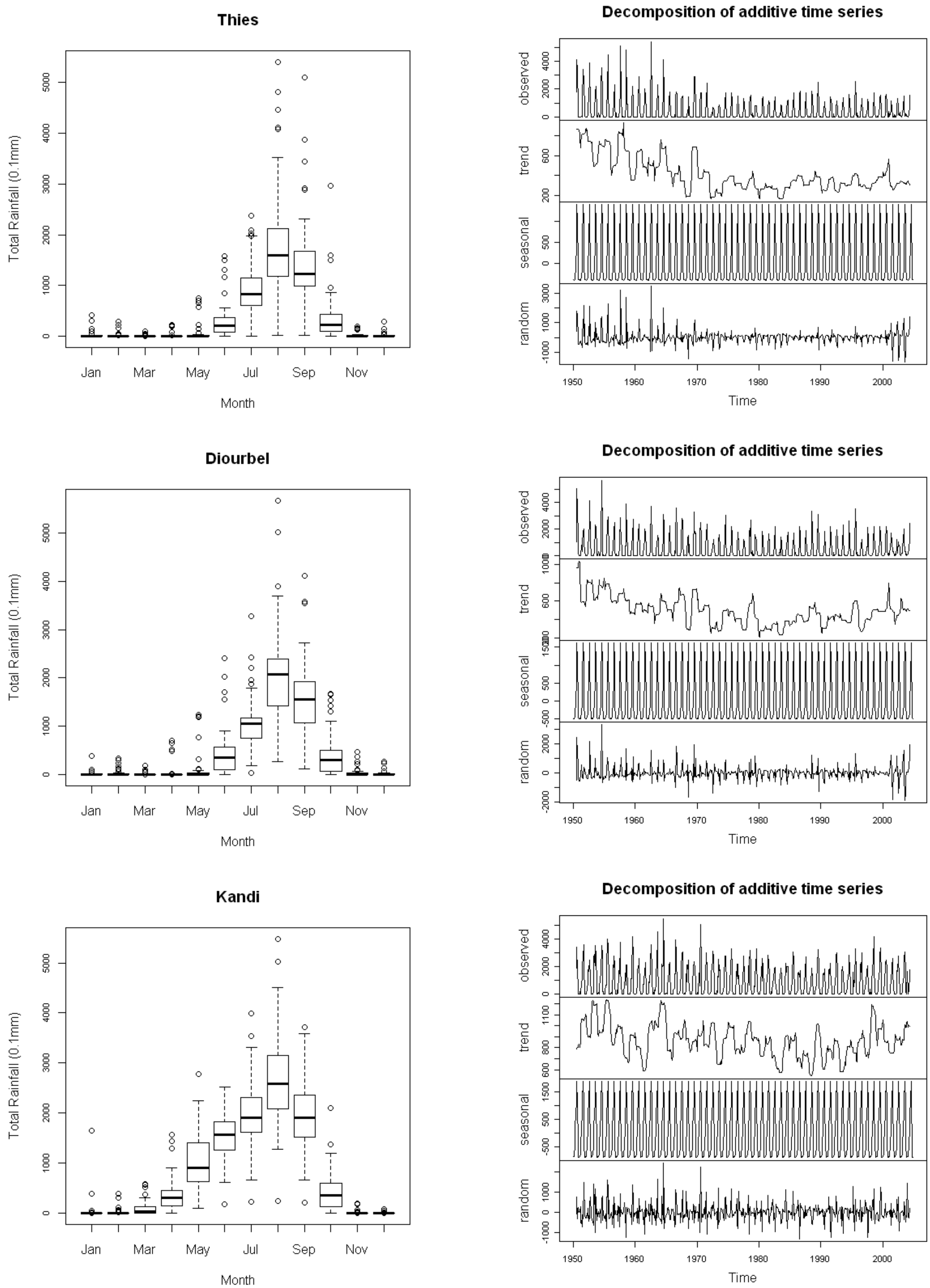


**Figure 2.** Contd.

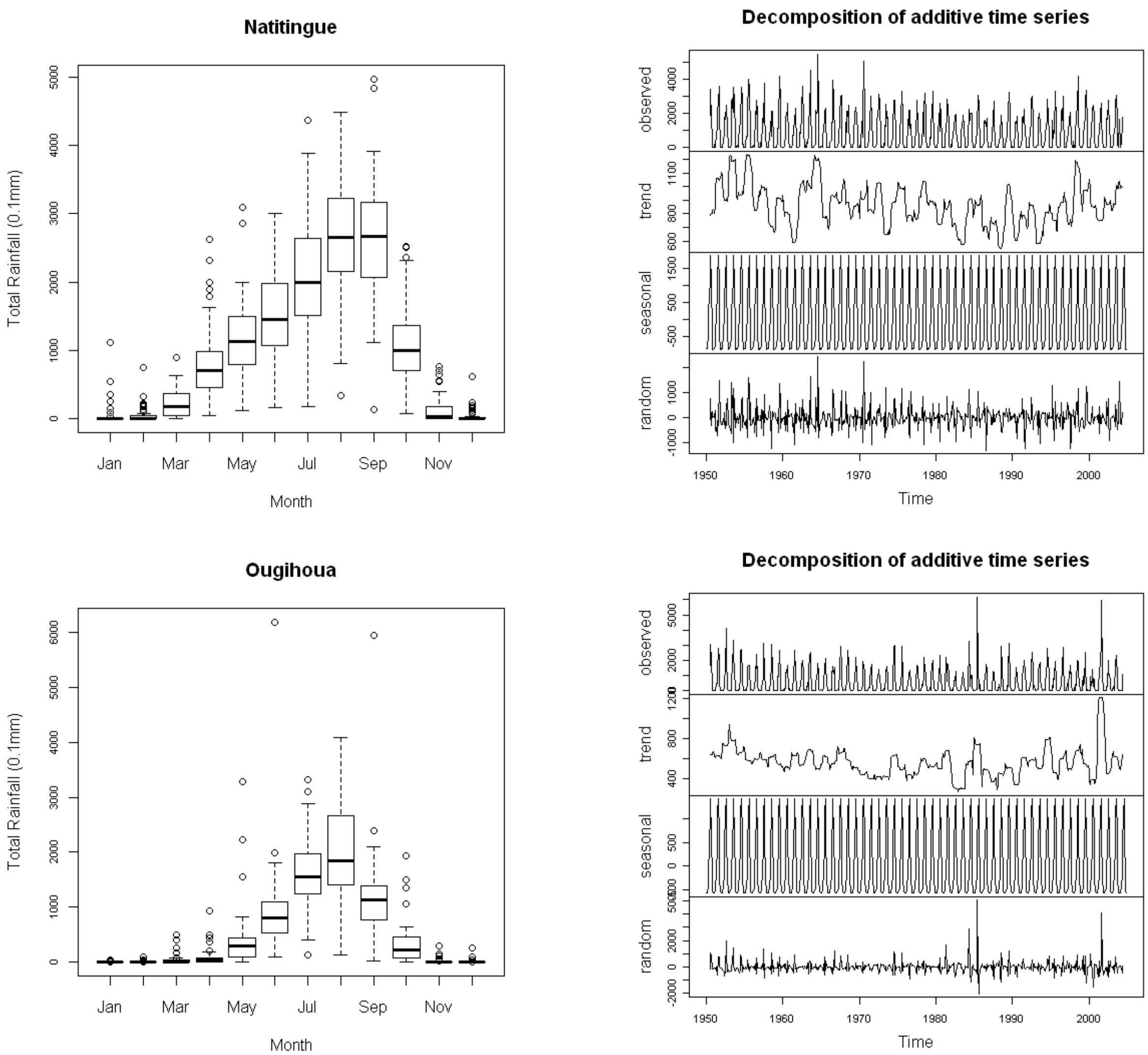


**Figure 2.** Contd.

pressure of the Azores through the North and the intensification of Santa Helena's high subtropical pressure, the trade winds of the southeast advect the wet air. This period is characterized by the penetration of the monsoon on the continent and the passage of the African Easterly Wave systems in direction to the Atlantic Ocean (Nicholson, 1993; Nicholson and Kim, 1997; Nicholson, 2000, 2005). These precipitation regimes are verified in all the analyzed stations.

## SEASONAL DISCRIMINANT ANALYSIS

The discriminant function analysis is used to determine which variables (rain-gauge stations) discriminate between two or more naturally occurring groups. Computationally, discriminant function analysis is very similar to (ANOVA). In practice, an observation is classified into a group if the squared distance (also called the Mahalanobis distance) of observation to the group center (mean) is the minimum.

For linear discriminant analysis (Table 2a) an assumption is made that covariance matrices are equal for all groups. There is a unique part of the squared distance formula for each group and that is called the linear discriminant function for that group. For any observation, the group with the smallest squared distance has the largest linear discriminant function and the observation is then classified into this group. The linear discriminant analysis has the property of symmetric squared distance: the linear discriminant function of group $i$ evaluated with the mean of group $j$ is equal to the linear discriminant function of group $j$ evaluated with the mean of group $i$.

For quadratic discriminant analysis (Table 2b) there is no assumption that the groups have equal covariance matrices. For linear discriminant analysis, an observation

**Table 2a.** The linear discriminant analysis summary of classification for the three seasons over the West Sahel Africa.

**Linear discriminant analysis:** Seasons versus Sahel rainfall
**Linear Method for Response:** SEASONS - 1 (WET): Jul-Sep; 2 (TRANS): Jun and Oct; 3 (DRY): Nov-May.
**Predictors:** TILLABERY; TAHOUA; MARADI; GAO; KITA; SEGOU; KOUTIALA; KIFFA; DAKAR; THIES; DIOURBEL; KANDI; NATITINGUE; OUGIHOUA

| Group | 1 | 2 | 3 |
|---|---|---|---|
| Count | 153 | 102 | 357 |

| | Summary of classification | | |
|---|---|---|---|
| **True group / Put into group** | **1** | **2** | **3** |
| **1** | 133 | 3 | 0 |
| **2** | 20 | 84 | 37 |
| **3** | 0 | 15 | 320 |
| Total N | 153 | 102 | 357 |
| N correct | 133 | 84 | 320 |
| Proportion | 0.869 | 0.824 | 0.896 |
| N = 612 | N correct = 537 | Proportion correct = 0.877 | |

**Table 2b.** The quadratic discriminant analysis summary of classification for the three seasons over the West Sahel Africa.

**Quadratic discriminant analysis: Seasons versus Sahel rainfall**
**Quadratic method for response:** SEASONS – 1 (WET): Jul-Sep; 2 (TRANS): Jun and Oct; 3 (DRY): Nov-May.
**Predictors:** TILLABERY; TAHOUA; MARADI; GAO; KITA; SEGOU; KOUTIALA; KIFFA; DAKAR; THIES; DIOURBEL; KANDI; NATITINGUE; OUGIHOUA

| Group | 1 | 2 | 3 |
|---|---|---|---|
| Count | 153 | 102 | 357 |

| | Summary of classification | | |
|---|---|---|---|
| **True group / Put into group** | **1** | **2** | **3** |
| **1** | 143 | 5 | 6 |
| **2** | 10 | 94 | 25 |
| **3** | 0 | 3 | 326 |
| Total N | 153 | 102 | 357 |
| N correct | 143 | 94 | 326 |
| Proportion | 0.935 | 0.922 | 0.913 |
| N = 612 | N correct = 563 | Proportion correct = 0.920 | |

is classified into the group that has the smallest squared distance. However, the squared distance does not simplify into a linear function. Unlike linear distance, quadratic distance is not symmetric. In other words, the quadratic discriminant function of group $i$ evaluated with the mean of group $j$ is not equal to the quadratic discriminant function of group $j$ evaluated with the mean of group $i$. On the results, quadratic distance is called the generalized squared distance (Johnson and Wichern, 1998).

In this manuscript, the discriminant analysis (Table 2a and b) was used to classify observations into three seasonal groups. The discriminant analysis can also be used to investigate how precipitation contributes to group

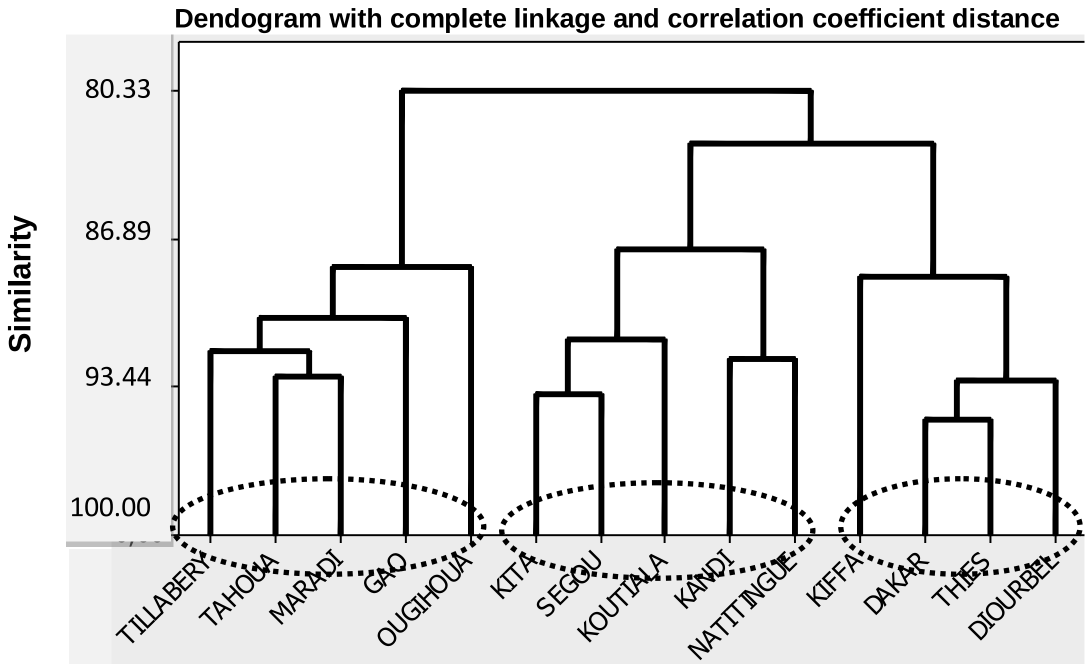


**Figure 3.** The dendrogram (Complete Linkage) displays the information in the amalgamation table in the form of a tree diagram, suggesting rain-gauge, which might be combined in three spatial groups, with total level of similarity (ρ) of about 88%.

separation by seasonality. As shown in the "summary of classification" in Table 2, the performed discriminant analysis correctly identified 92% (quadratic discriminant analysis) of the weather variability classification, considering the available data of the 14 stations that provide continuous month records of precipitation.

## Spatial cluster analysis

The Sahel is represented by the leading pattern of observed rainfall variability, capturing at least 87% of the total variance (Figure 3). The clustering of variables was used to classify variables into groups when the groups are initially not known. One reason to cluster variables may be to classify homogeneous regions in the West Sahel Africa. The final grouping of clusters (also called the final partition) is the grouping of clusters that will, hopefully, identify groups whose observations or variables share common characteristics.

The dendrogram with complete linkage (Figure 3) suggest variables which might be combined in three groups, with total level of similarity ( $\rho$ ) of about 88%, classifying similar rainfall regimes (defining similar spatio-temporal rainfall regimes). The final partition is given by: (R1) Tillabery, Tahoua, Maradi, Gao and Ougihoua ($p$ = 89%); (R2) Kita, Segou, Koutiala, Kandi and Natitingue ( $\rho$ = 88%); (R3) Kiffa, Dakar, Thies and Diourbel $\rho$ = 87%). Further work could explore these connections, verifying if monthly total precipitation in Mali (Senegal) tends to indicate precipitation in monthly total Burkina Faso (Mauritania), for instance. These variables could contain similar information and be accurately combined.

The discriminant analysis was also used to investigate how precipitation contributes to spatial network partition. As shown in the "summary of classification" in Table 3, the discriminant analysis identified the seasonal variability of the spatial network, with probability of well-classification grater than 82%, classifying similar rainfall regimes: (R1) Tillabery, Tahoua, Maradi, Gao, and Ougihoua; (R2) Kita, Segou, Koutiala, Kandi, Natitingue and (R3) Kiffa, Dakar, Thies, Diourbel.

## LOWESS SMOOTHER TREND DETECTION

Figure 4 illustrates the rainfall trend for each spatial classification (R1, R2 and R3), using the time series plot with a LOWESS smoother for each of three groups. The LOWESS smoother (a computational intensive method) is a line that is fitted to the data to explore the potential relationships between two variables, in this case the temporal dynamics of annual total rainfall, without fitting a specific model, such as a linear regression with a theoretical distribution. In practice, LOWESS, originally proposed by Cleveland (1979) and further developed by Cleveland and Devlin (1988), specifically denotes a method descriptively known as locally weighted polynomial

**Table 3.** The quadratic discriminant analysis summary of classification for the three spatial networks over the West Sahel Africa.

**Discriminant analysis: Seasons versus Region R1**

**Quadratic method for response:** SEASONS – 1 (WET): Jul-Sep; 2 (TRANS): Jun and Oct; 3 (DRY): Nov-May.
**Predictors:** TILLABERY; TAHOUA; MARADI; GAO; OUGIHOUA

| Group | 1 | 2 | 3 |
|---|---|---|---|
| Count | 153 | 102 | 357 |

| | Summary of classification | | |
|---|---|---|---|
| Put into group \ True group | **1** | **2** | **3** |
| **1** | 119 | 9 | 5 |
| **2** | 34 | 53 | 21 |
| **3** | 0 | 40 | 331 |
| Total N | 153 | 102 | 357 |
| N correct | 119 | 53 | 331 |
| Proportion | 0.778 | 0.520 | 0.927 |
| N = 612 | N correct = 503 | Proportion correct = 0.822 | |

**Discriminant analysis: Seasons versus Region R2**

**Quadratic method for response:** SEASONS – 1 (WET): Jul-Sep; 2 (TRANS): Jun and Oct; 3 (DRY): Nov-May.
**Predictors**: KITA; SEGOU; KOUTIALA; KANDI; NATITINGUE

| Group | 1 | 2 | 3 |
|---|---|---|---|
| Count | 153 | 102 | 357 |

| | Summary of classification | | |
|---|---|---|---|
| Put into group \ True group | **1** | **2** | **3** |
| **1** | 134 | 10 | 8 |
| **2** | 19 | 81 | 36 |
| **3** | 0 | 11 | 313 |
| Total N | 153 | 102 | 357 |
| N correct | 134 | 81 | 313 |
| Proportion | 0.876 | 0.794 | 0.877 |
| N = 612 | N correct = 528 | Proportion correct = 0.863 | |

**Discriminant analysis: Seasons versus Region R3**

**Quadratic method for response:** SEASONS – 1 (WET): Jul-Sep; 2 (TRANS): Jun and Oct; 3 (DRY): Nov-May.
**Predictors:** KIFFA; DAKAR; THIES; DIOURBEL

| Group | 1 | 2 | 3 |
|---|---|---|---|
| Count | 153 | 102 | 357 |

| | Summary of classification | | |
|---|---|---|---|
| Put into group \ True group | **1** | **2** | **3** |
| **1** | 123 | 8 | 0 |
| **2** | 30 | 80 | 18 |
| **3** | 0 | 14 | 339 |
| Total N | 153 | 102 | 357 |
| N correct | 123 | 80 | 339 |
| Proportion | 0.804 | 0.784 | 0.950 |
| N = 612 | N correct = 542 | Proportion correct = 0.886 | |

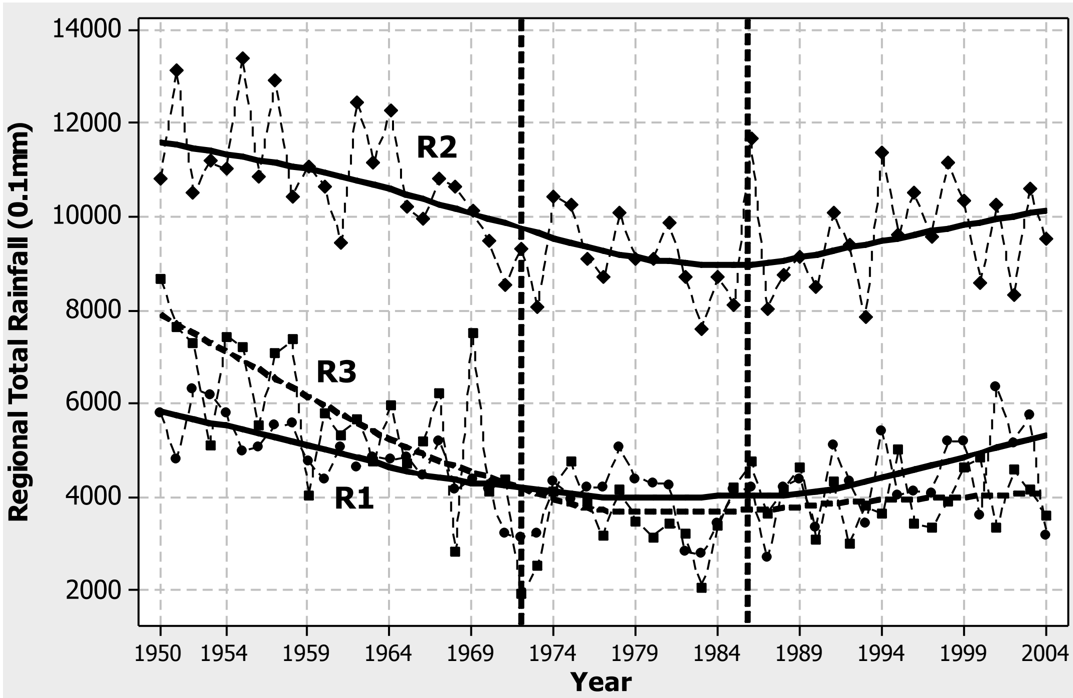


**Figure 4.** The trend analysis of the annual average rainfall for each one of the 3 spatiotemporal clusters on the West Sahel Africa (1950 to 2000): R1, R2 and R3. Structural change-points can be easily observed of about 1972, and 1986, which can be associated to the PDO index. The continuous line (as a layout device) fits a Robust LOWESS (Cleveland, W. S., 1979: Robust locally weighted regression and smoothing scatter plots. *Journal of the American Statistical Association*, **74**: 829-836.) curve to the annual mean rainfall (1950-2000) for the meteorological stations in each spatial network cell: Eastern West Sahel Africa; Southern West Sahel Africa and Extreme West Sahel Africa.

regression. The LOWESS fit produce a function that captures the deterministic structure in the data quite well.
The decreasing trend of above-average rainfall from 1950's to early 1970's and the increasing trend of below-average rainfall since the early of the 1980's are perceptible. The conditions in recent years seem that rainfall has increased to levels similar to those in the 1950s. The fluctuations between "wet" and "dry" in the Sahel/Soudan zones are extreme even on decadal and multi-decadal time scales. The comparatively wet conditions of the 1950s are unlike any that occurred since late in the 19th century (Nicholson, 1993, 1996). Those of the dry episode from roughly 1968 to 1997 are comparable only to those that prevailed in the 1820s and 1830s, an arid period that likewise affected nearly all of Africa[2]. Thus, it is of interest to take a more detailed look at how the recent 6 years compare with these episodes. In this section the mean rainfall in the various sectors, the seasonal cycle and the location of key isohyets are examined for three periods: the "wet" decade of 1950 to 1959, the "dry" decades of 1968 to 1997, and the recent 7 years 1998-2004.

## PRINCIPAL COMPONENT ANALYSIS

Concerning the West Sahel July to September season (Figure 5) the first principal component has variance (eigenvalue) 5.85 and accounts for 41.8% of the total variance. The second principal component has variance 2.04 and accounts for 14.6% of the data variability. Together, the first two principal components represent (capture) 56.4% of the total variability. One could think of the first principal component as representing an overall spatial rainfall pattern, because the coefficients of these terms have the same sign and are not close to zero. The second principal component could be thought of as contrasting annual rainfall variability. Regarding the West Sahel June and October season (Figure 6) the first principal component has variance (eigenvalue) 5.47 and

[2] Around 1790 dry conditions similar to those of the late 20th century set in (Rain, 1999) and continued until around 1870. After that, a very wet period set in for around 25 years, followed by a return to drier conditions. While the drying begun around 1895 and caused its first large famine only in the early 20th century, the 1820s and 1830s saw a 12 to 15 year drought and regional instances of major famine from Senegal to Chad. Historical records suggest this drought caused a large-scale emigration from the Bornu Empire, contributing to its rapid decline in the 19th century (Lovejoy and Baier, 1975). In what is now northern Senegal, the Kingdom of Fouta Tooro was struck by a famine caused by the failure of 1833's rainy season, leading to waves of famine until 1837 (Curtin, 1975).

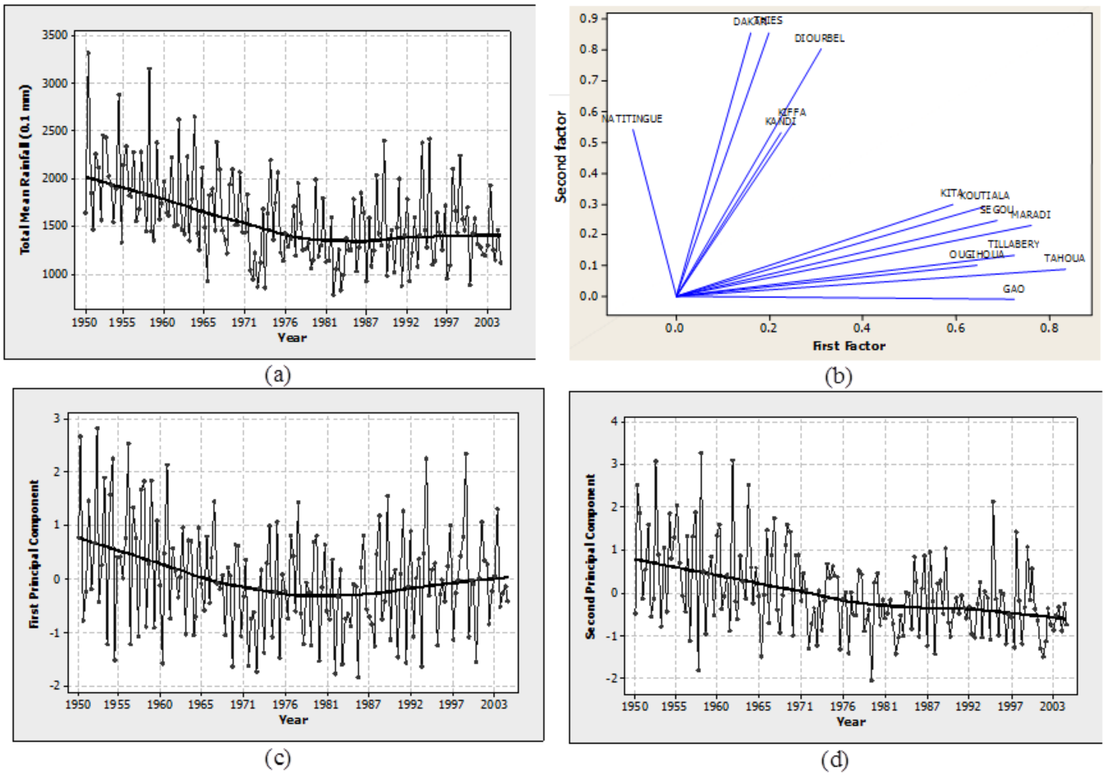

**Figure 5.** West Sahelian JAS season - (a) The time series plot of the Total Rainfall ; (b) The loading plot; (c) The time series plot of the PC1 ; (d) The time series plot of the PC2.

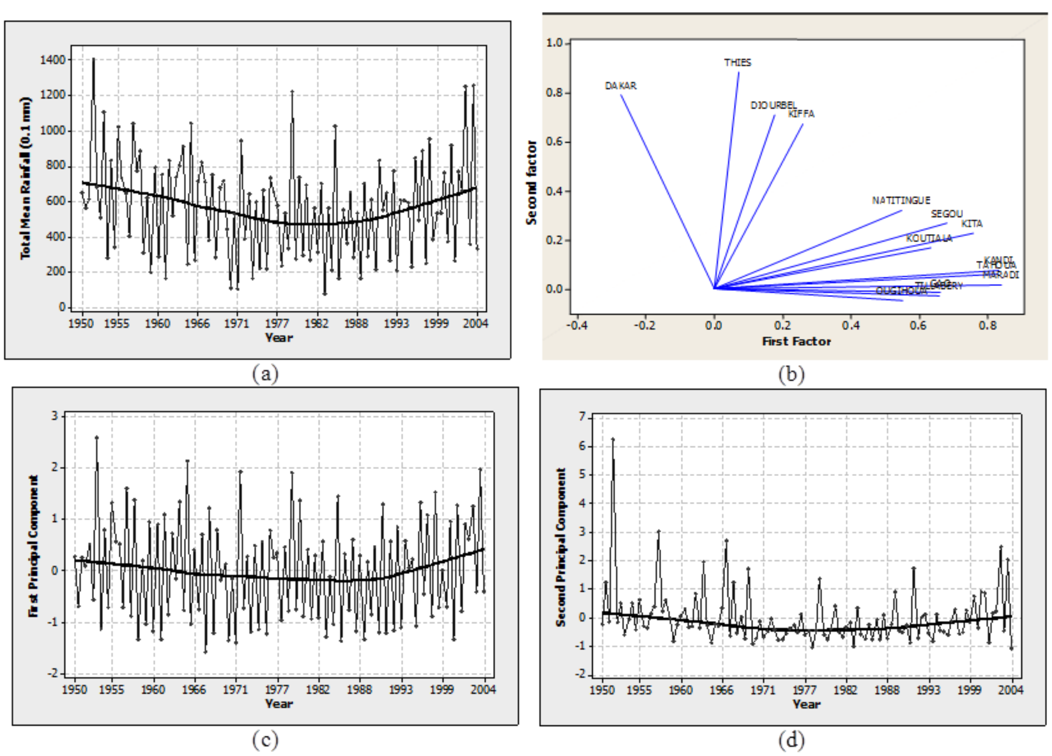

**Figure 6.** West Sahelian J&O season - (a) The time series plot of the Total Rainfall ; (b) The loading plot; (c) The time series plot of the PC1 ; (d) The time series plot of the PC2.

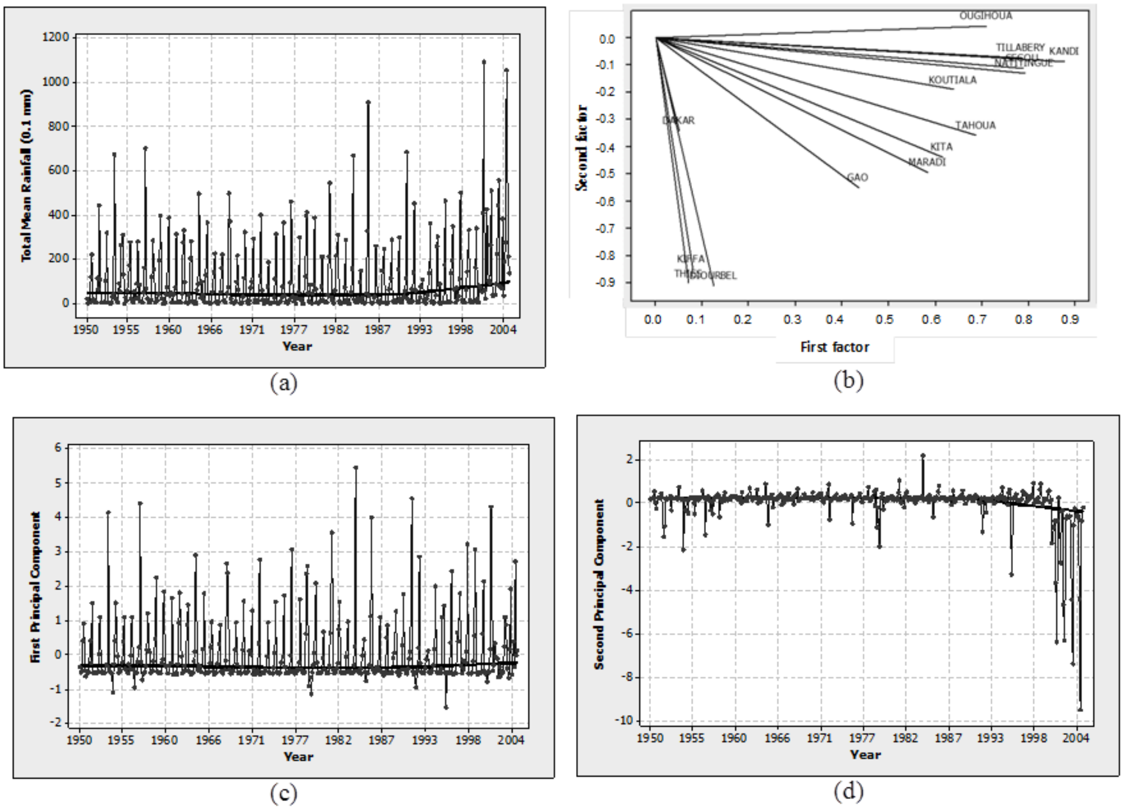


**Figure 7.** West Sahelian NDJFMAM season (a) The time series plot of the Total Rainfall ; (b) The loading plot; (c) The time series plot of the PC1 ; (d) The time series plot of the PC2.

accounts for 39.1% of the total variance. The second principal component has variance 2.63 and accounts for 18.8% of the data variability. Together, the first two principal components correspond to 57.9% of the total variability. For the West Sahel November to May season (Figure 7) the first principal component has variance (eigenvalue) 5.60 and accounts for 40.0% of the total variance. The second principal component has variance 2.43 and accounts for 17.4% of the data variability.

Together, the first two principal components represent 57.4% of the total variability. In these two cases, the second principal component could be thought of as contrasting spatial annual rainfall variability, sorting out Kiffa, Dakar, Thies and Diourbel from the other rain-gauge stations.

In this study, the second principal component, under a meteorological point of view, is the more important, since the distinct spatial structures intend to explain the anomalies linked to the dynamical variability of the ITCZ.

## PRINCIPAL COMPONENTS REGRESSION

In statistics, principal component regression (PCR) is a regression analysis that uses principal component analysis when estimating regression coefficients. It is a procedure used to overcome problems which arise when the exploratory variables are close to being collinear. The use of PCR has received a lot of attention in the literature in the past few years (Jolliffe, 1982), and the topic is now beginning to appear in textbooks. The PCR performs simple and multiple regression using least squares.

In Table 4 the *p-value* in the (ANOVA) table shows that the model estimated by the regression procedure is significant at a $\alpha = 0.05$ level. This indicates that at least one coefficient is different from zero. The $R^2$ value indicates that the predictors capture 97.3% of the variance in PC1. Proceeding, the residual analysis, one can identify the observation 25 (July 1958) as unusual Sahel Total Rainfall for July to September Season, with a large standardized residual; this may indicate it is outlier. In Figure 8 the histogram indicates that outliers may not exist in the data, the normal probability plot shows an approximately linear pattern consistent with a normal distribution. The plot of residuals versus the fitted values may indicate the residuals have constant variance. In Table 5, also, the *p-value* in the (ANOVA) table shows that the model estimated by the regression procedure is significant at a $\alpha = 0.05$ level. This indicates that at least one coefficient is different from zero. The $R^2$ value indicates that the predictors capture 90.4% of the variance

**Table 4.** JAS Multiple and Simple Regression using least squares. Use this procedure for fitting general least squares models, storing regression statistics, examining residual diagnostics, generating point estimates, generating prediction and confidence intervals, and performing lack-of-fit tests. (Top) Regression Analysis: The Sahel JAS Season TOTAL RAIN versus PC1 and PC2; (Bottom) Regression Analysis: The Sahel JAS Season TOTAL RAIN versus PC1.

**The regression equation is total rain = 22,354 – 2,733*PC1 – 683*PC2**

| Predictor | Coeff. | SE Coeff. | T | P |
|---|---|---|---|---|
| Constant | 22354.3 | 41.2 | 542.40 | 0.000 |
| PC1 | -2733.16 | 17.09 | -159.93 | 0.000 |
| PC2 | -683.00 | 28.97 | -23.58 | 0.000 |
| S = 509.783 | R-Sq = 99.4% | | R-Sq(adj) = 99.4% | |

| Analysis of variance / Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| Regression | 2 | 6791830032 | 3395915016 | 13067.33 | 0.000 |
| Residual Error | 150 | 38981747 | 259878 | | |
| Total | 152 | 6830811779 | | | |

**The regression equation is total rain = 22354 – 2733*PC1**

| Predictor | Coef | SE Coef | T | P |
|---|---|---|---|---|
| Constant | 22354.3 | 89.1 | 250.86 | 0.000 |
| PC1 | -2733.16 | 36.95 | -73.97 | 0.000 |
| S = 1102.24 | R-Sq = 97.3% | | R-Sq(adj) = 97.3% | |

| Analysis of variance / Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| Regression | 1 | 6647355969 | 6647355969 | 5471.35 | 0.000 |
| Residual Error | 151 | 183455811 | 1214939 | | |
| Total | 152 | 6830811779 | | | |

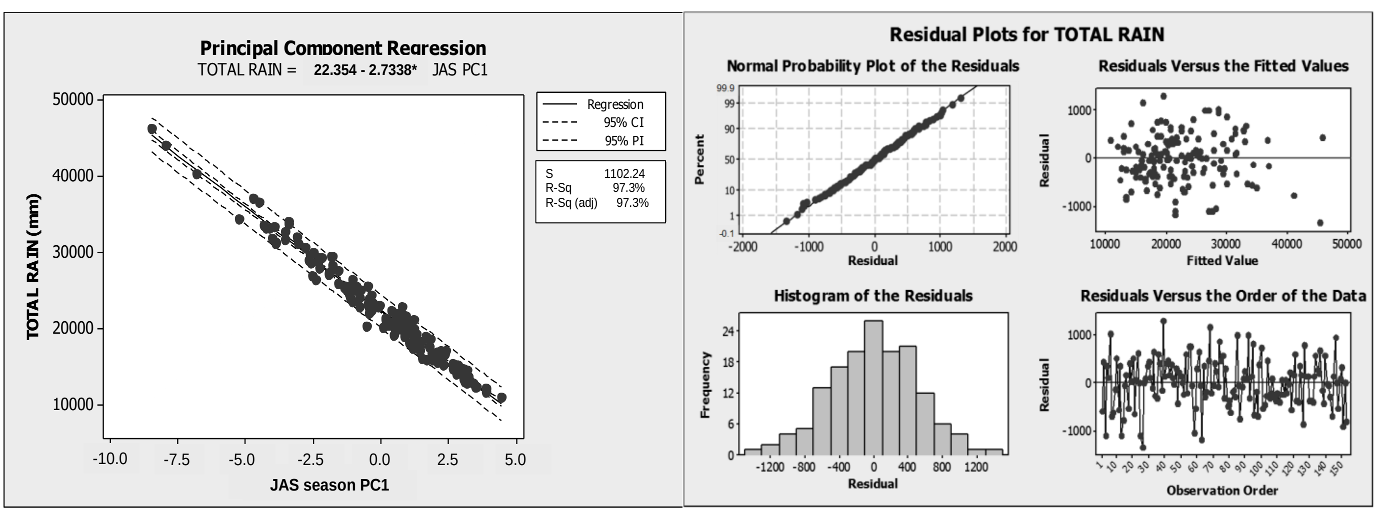


**Figure 8.** West Sahelian JAS season (a) The PCR fitted line plot ; (b) The PCR residual diagnostics.

in PC1. The observations 4, 16, 34 and 84 (October 1951, October 1957, October 1966, October 1991, respectively) are identified as unusual Sahel Total Rainfall for July and October Season, with a large standardized residual; this may indicate they are outliers. In Figure 9 the histogram indicates that outliers may exist in the data, the normal probability plot shows an approximately linear pattern consistent with a normal distribution.

**Table 5.** J&O Multiple and Simple Regression using least squares. Use this procedure for fitting general least squares models, storing regression statistics, examining residual diagnostics, generating point estimates, generating prediction and confidence intervals, and performing lack-of-fit tests. (Top) Regression Analysis: The Sahel J&O Season TOTAL RAIN versus PC1 and PC2; (Bottom) Regression Analysis: The Sahel J&O Season TOTAL RAIN versus PC1.

**The regression equation is total rain = 7817 - 1571*PC1 + 668*PC2**

| Predictor | Coeff. | SE Coeff. | T | P |
|---|---|---|---|---|
| Constant | 7817.39 | 49.54 | 157.79 | 0.000 |
| PC1 | 1570.62 | 21.29 | -73.78 | 0.000 |
| PC2 | 667.92 | 30.68 | 21.77 | 0.000 |
| S = 500.366 | R-Sq = 98.4% | | R-Sq(adj) = 98.3% | |

| Source \ Analysis of variance | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| Regression | 2 | 1481652617 | 740826308 | 2958.97 | 0.000 |
| Residual Error | 99 | 24786220 | 250366 | | |
| Total | 101 | 1506438837 | | | |

**The regression equation is total rain = 7817 - 1571*PC1**

| Predictor | Coeff. | SE Coeff. | T | P |
|---|---|---|---|---|
| Constant | 7817.4 | 118.6 | 65.93 | 0.000 |
| PC1 | -1570.62 | 50.95 | -30.83 | 0.000 |
| S = 1197.58 | R-Sq = 90.5% | | R-Sq(adj) = 90.4% | |

| Source \ Analysis of variance | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| Regression | 1 | 1363020010 | 1363020010 | 950.38 | 0.000 |
| Residual Error | 100 | 143418827 | 1434188 | | |
| Total | 101 | 1506438837 | | | |

The point 4 (October 1951) in the upper-right corner of the plot may be outlier, the same point that is labeled unusual observations in the output.

In Table 6, as well, the *p-value* in the (ANOVA) table shows that the model estimated by the regression procedure is significant at a $\alpha = 0.05$ level. This indicates that at least one coefficient is different from zero. The $R^2$ value indicates that the predictors capture 97.3% of the variance in PC1. In Figure 10 the histogram indicates that outliers may exist in the data, the normal probability plot shows a non linear pattern.

Data with large standardized residuals, identified as unusual in the Sahel rainfall data sets, may be thought to characterize events of the incursion and persistence, or stationarity, of ITCZ over the region rather than being outliers.

## FINAL REMARKS

This work provides a report of worthwhile results that variability of rainfall in the Sahel results from the temporal variability of the southern incursions of ITCZ on West Africa. It is significant in the differences in the rainfall regimes between three succeeding periods 1950 to 1969 a negative trend, 1970 to 1984 there is no statistically significant trend and 1985 to 2004 a positive tendency. The interannual rainfall discrepancy between the wet and the dry periods is 180 mm/yr. This difference is relatively evenly distributed in space, with no clear meridional gradient. Between these two periods, the parameter measure of the occurrence rate displays a systematic decrease, which appears well correlated to the decrease of the mean interannual rainfall. When looking at the intraseasonal scale, it appears that the rainfall deficit of the dry period is primarily linked to a deficit of the number of events occurring during the core of the rainy season over the Sahel, and during the first rainy season for the region extending south 9 to10°N. It is also shown that, in the south, the dry period is characterized by a shift in time of the second rainy season. All these characteristics have strong implications in term of agricultural and water resources management. They also raise questions about the traditional scheme used to characterize the dynamics of the West African monsoon.

Desertification is still often viewed as an irreversible process triggered by a combination of declining rainfall and destructive farming methods. There is also a strong confusion over why the Sahel is becoming green.

Scientists believe the main reason is increased rainfall

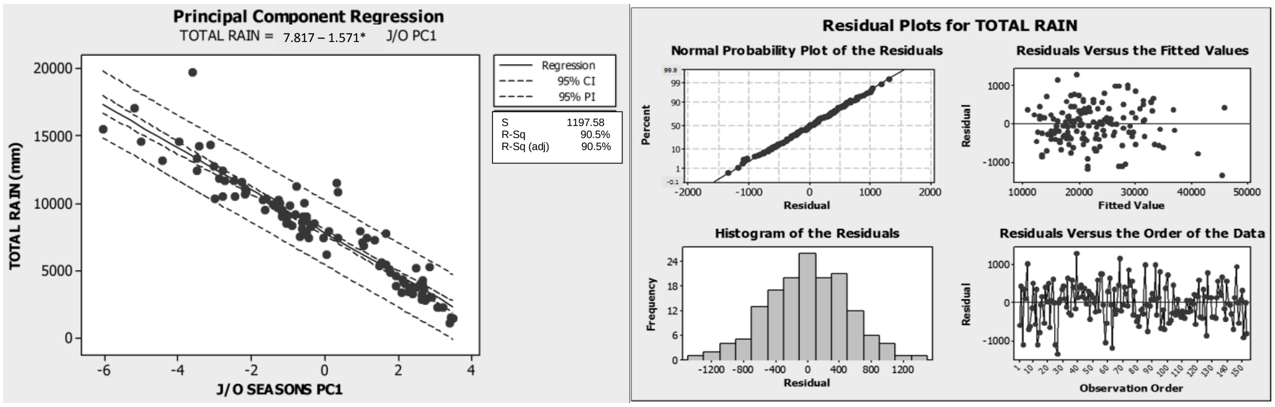


**Figure 9.** West Sahelian J&O season (a) The PCR fitted line plot; (b) The PCR residual diagnostics.

**Table 6.** NDJFMAM Multiple and Simple Regression using least squares. Use this procedure for fitting general least squares models, storing regression statistics, examining residual diagnostics, generating point estimates, generating prediction and confidence intervals, and performing lack-of-fit tests. (Top) Regression Analysis: The Sahel N-M Season TOTAL RAIN versus PC1 and PC2; (Bottom) Regression Analysis: The Sahel N-M Season TOTAL RAIN versus PC1.

| **The regression equation is total rain = 1208 - 809*PC1 + 59,1*PC2** | | | | |
|---|---|---|---|---|
| **Predictor** | **Coeff.** | **SE Coeff.** | **T** | **P** |
| Constant | 1207.52 | 23.32 | 51.79 | 0.000 |
| PC1 | -808.931 | 9.865 | -82.00 | 0.000 |
| PC2 | 59.06 | 14.98 | 3.94 | 0.000 |
| S = 440.550 | R-Sq = 95.0% | | R-Sq(adj) = 95.0% | |

| Source / Analysis of variance | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| Regression | 2 | 1307949333 | 653974667 | 3369.53 | 0.000 |
| Residual Error | 354 | 68705992 | 194085 | | |
| Total | 356 | 1376655325 | | | |

| **The regression equation is total rain = 1208 - 809*PC1** | | | | |
|---|---|---|---|---|
| **Predictor** | **Coeff.** | **SE Coeff.** | **T** | **P** |
| Constant | 1207.52 | 23.79 | 50.76 | 0.000 |
| PC1 | -808.93 | 10.07 | -80.37 | 0.000 |
| PC2 | 59.06 | 14.98 | 3.94 | 0.000 |
| S = 449.486 | R-Sq = 94.8% | | R-Sq(adj) = 94.8% | |

| Source / Analysis of variance | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| Regression | 1 | 1304931944 | 1304931944 | 6458,85 | 0.000 |
| Residual Error | 355 | 71723380 | 202038 | | |
| Total | 356 | 1376655325 | | | |

since the great droughts of the early 1970's and 1980's. But farmers have also been adopting better methods of keeping soil and water on their land. Besides ENSO (*El Nino* Southern Oscillation), the NAO(North Atlantic Oscillation), and west African climate anomaly patterns, other continental-scale and sub continental climate anomalies play significant roles in determining inter-annual and longer climate variability time scales (Nicholson, 2000, 2005). The decades 1950 to 1969 was characterized by above-normal precipitation over the West Sahel region. Later, during the period 1970 to 1984, this rainfall anomaly pattern dramatically reversed in sign, with rainfall deficits observed for most of Sahel; these two time periods are the reversal in the sign of the Sahelian

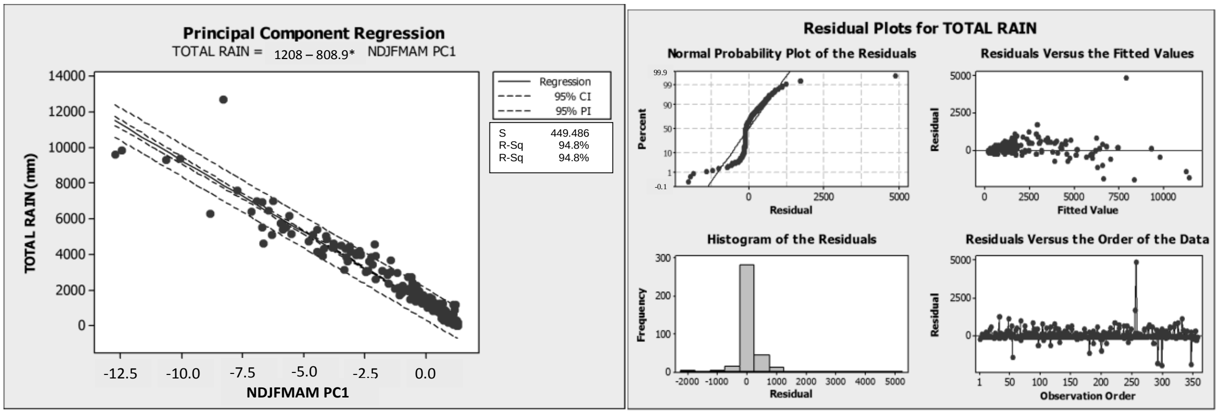


**Figure 10.** West Sahelian NDJFMAM season (a) The PCR fitted line plot ; (b) The PCR residual diagnostics.

rainfall anomalies. More recently, the pattern has been one of decreased aridity throughout most of the Sahelian region; mean rainfall increased around 10% (the angular coefficient in trend analysis) in the Sahel region. In comparison with the period between 1950 and 1970, the average length of the rainy season has not changed significantly during the dry period 1970 to1984. Instead, the decrease in rainfall in July and August explains most of the diminution of total annual rainfall over the Sahel since 1970.

## ACKNOWLEDGEMENTS

The authors would like to thank the reviewers for their comments that help to improve this manuscript. P. S. Lucio was sponsored by a POCTI grant (SFRH/BPD/5614/2001) from FCT (Portugal), and at the present time he is sponsored by a PQ2 grant (302493/2007-7) from CNPq (Brazil).

### REFERENCES

Agnew CT, Chappell A (2000). Drought in the Sahel. Geo. J., 48: 299-311.

Charney JG, Stone PH, Quirk WJ (1975). Drought in the Sahara: A biogeophysical feedback mechanism. Sci. 187: 434-435.

Charney JG (1975). Dynamics of deserts and drought in Sahel. Quart. J. Royal Meteor. Soc., 101: 193-202.

Cleveland WS (1979). Robust locally weighted regression and smoothing scatterplots. J. Am. Statistical Assoc., 74: 829-836.

Cleveland WS, Devlin SJ (1988). Locally-Weighted Regression: An Approach to Regression Analysis by Local Fitting. J. Am. Statistical Assoc., 83 (402): 596–610.

Folland CK, Parker DE, Palmer TN (1986). Sahel rainfall and worldwide sea temperatures 1901-85. Nature, 320: 602-607.

Giannini A, Saravanan R, Chang P (2003). Ocean forcing of Sahel rainfall on interannual to interdecadal time scales. Science, 302: 1027-1030.

Harrison SP, Jolly D, Laarif F, Abe-Ouchi A, Dong B, Herterich K, Hewitt C, Joussaume S, Kutzbach JE, Mitchell J, de Noblet N, Valdes P (1998). Intercomparison of simulated global vegetation distribution in response to 6kyr B.P. orbital forcing. J. Clim., 11: 2721-2742.

Hulme M (1996). Recent climatic change in the world's drylands. Geophys. Res. Lett. 23: 61-64.

Johnson RA, Wichern DW (1998). Applied multivariate statistical analysis. 4th ed., New Jersey: Prentice Hall, New Jersey, p. 616.

Jolliffe IT (1982). A note on the Use of Principal Components in Regression. J. R. Stat. Soc. Series C (Appl. Stat.) 31(3): 300–303.

Milich L (1997). Expansion and contraction of the Sahara desert. Science, 253: 299-301.

Nicholson S (1993). An overview of African rainfall fluctuations of the last decade. J. Clim., 6: 1463–1466.

Nicholson S (1996). Environmental Change within the Historical Period. In: AS Goudie, WM Adams, A Orme, Editors, *The Physical Geography of Africa*, Oxford University Press, Oxford, pp. 60–75.

Nicholson S, Kim J (1997). The relationship of the El Niño Southern Oscillation to African rainfall. Int. J. Climatolo., 17: 117–135.

Nicholson S (2000). The nature of rainfall variability over Africa on time scales of decades to millennia. Glob. Planet. Change. 26: 137–158.

Nicholson S (2005). On the question of the 'recovery' of the rains in the West African Sahel. J. Arid. Environ., 63: 615–641.

R Development Core Team (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rubin DB (1987). Multiple imputation for nonresponse in Surveys. New York: John Wiley and Sons.

Rubin DB (1996). Multiple imputation after 18+ years (with discussion). J. Am. Statist. Assoc., 91: 473-489.